9-9-05

Australian Government

Patent Office
Canberra

I, LEANNE MYNOTT, MANAGER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2002950183 for a patent by THE COUNCIL OF THE QUEENSLAND INSTITUTE OF MEDICAL RESEARCH as filed on 12 July 2002.

WITNESS my hand this
Thirtieth day of August 2005

LEANNE MYNOTT
MANAGER EXAMINATION SUPPORT
AND SALES

AUSTRALIA

Patents Act 1990

The Council of the Queensland Institute of Medical Research

PROVISIONAL SPECIFICATION

Invention Title:

Expression of hydrophobic proteins

The invention is described in the following statement:

Field of the Invention:

This invention relates to the expression of non-native (ie heterologous) polypeptides which comprise a proportion of hydrophobic amino acids in an expression system such as a bacterial host (eg *E. coli*). In particular, the invention provides a method for designing
5   candidate polyepitope polypeptides with an increased probability of being efficiently expressed (ie in amounts detectable by SDS-PAGE). One particular application of the invention relates to the production of a polyepitope polypeptide comprising 26 CTL epitopes from Epstein-Barr virus (EBV) for use in a polyepitope vaccine capable of eliciting a CTL immune response for the prevention of diseases associated with EBV.

10

Background of the Invention:

In order to maximise production of recombinant polypeptides in a bacterial host (eg *E. coli*), a number of parameters can be considered including factors affecting transcription (eg promoter choice, etc) and factors affecting translation mechanisms such as minimising
15   the use of rare codons. However, these are unlikely to have an impact on the production of recombinant polypeptides comprising stretches of hydrophobic amino acids, which have traditionally proven difficult to produce in recombinant bacterial expression systems. Indeed, in the case of polypeptides comprising transmembrane sequences, the removal of these hydrophobic sequences generally improves yields of the recombinant molecule
20   (Frace et al, 1999; Hobman et al, 1994; Polte et al, 1991; EMBL website- protein toxicity: www.embl-heidelberg.de/ExternalInfo).

The most likely reason for problems occurring in the production of foreign polypeptides possessing regions of hydrophobicity (particularly those with non-native sequences such as fusion proteins), is the post-translational association of nascent
25   polypeptides with chaperone proteins (eg *E. coli* groEL). GroEL is involved in the refolding process of polypeptides emerging from the ribosome and proteins will recycle through the chaperone system until the correct conformation is achieved or the protein is targeted for degradation. GroEL is known to bind hydrophobic amino acids and part of the refolding process is essentially to bury these hydrophobic sequences within the interior
30   of the protein (Fisher and Yuan, 1994; Zahn and Pluckthun, 1994; Hayer-Hartl et al, 1994; Richarme and Kohiyama, 1994; Hendrick and Hartl, 1995; Lin et al, 1995).

Polyepitope or "polytope" constructs (ie polypeptides comprising a tandem array of epitopes which may be contiguous or otherwise spaced apart by short intervening amino acid sequences of, for example, 1 to 5 amino acids in length), would be expected to be

inherently unable to internalise any hydrophobic regions as they are not naturally-occurring sequences and lack the folding capabilities inbuilt in native proteins. Hence polypeptides which consist of non-native sequences, particularly those with a high proportion of hydrophobic amino acids, are likely to be sequestered in the chaperone
5   folding pathway and ultimately targeted for degradation if a certain degree of conformational stability cannot be achieved.

Polyepitope vaccines typically comprise one or more polypeptides each made up of a tandem array of amino acid CTL epitopes. These CTL epitopes, particularly those of the HLA A2 type, often comprise predominantly hydrophobic amino acids and since HLA A2
10   is represented in over 40% of the human population it is mandatory that these epitopes be included in any effective polyepitope vaccine formulation. Examples of such polyepitope vaccines are described in Australian Patent No. 736336, the entire disclosure of which is to be regarded as being incorporated herein by reference. In this patent, vaccines are described which comprise a synthetic or recombinant polypeptide, or a recombinant
15   vaccinia virus or DNA vaccine encoding same, wherein the synthetic or recombinant polypeptide typically comprises a tandem array of CTL epitopes (eg 2 to 10) wherein at least two of the CTL epitopes are contiguous or spaced apart by intervening sequences in which the intervening sequences do not (i) comprise methionine or (ii) comprise naturally occurring flanking sequences of the epitopes. Particularly described in the prior patent are
20   vaccines comprising a polyepitope vaccinia virus encoding a polyepitope polypeptide comprising 9 CTL epitopes (each of 9 to 10 amino acids in length) from EBV. Standard chromium release assays conducted with this virus in a panel of target cells expressing the HLA alleles for restriction of each epitope and using autologous CTL clones specific for each epitope as effector cells, showed that each epitope could be efficiently processed from
25   the polyepitope polypeptide since, in all cases, the CTL clones recognised and killed the HLA matched target cell infected with the polyepitope vaccinia virus, but did not kill any of the negative controls (ie TK-vaccinia).

Polyepitope vaccines may include a large number of CTL epitopes (eg 10 or more) so that the HLA diversity of the target population is covered. It has therefore been
30   contemplated by the present applicant to produce an EBV polyepitope vaccine which includes EBV epitopes restricted by HLA A2, A3, A11, A23, A24, B7, B8, B27, B35, B44, B46, B57, B58, B60 and B62, so as to provide protection against EBV in over 90% of the human population. This would involve the incorporation of about 26 EBV CTL epitopes into a polyepitope polypeptide. For the reasons given above, it was expected that such a

4

polypeptide would contain hydrophobic regions and that expression in a bacterial host would be highly problematical.

The work leading to the present invention was aimed at elucidating a method or procedure for overcoming the difficulties of expressing non-native, polypeptides which comprise a proportion of hydrophobic amino acids (eg polyepitope polypeptides) in a bacterial host such as *E. coli*. The present applicants have, as a result of that work, identified a novel method for designing candidate polyepitope polypeptides with an increased probability of being efficiently expressed in a bacterial host and/or yielding a purified polyepitope polypeptide which is soluble in aqueous solutions. The method involves identifying one or more hydrophobic peptide sequences within a polypeptide and arranging or re-locating at least one of the hydrophobic peptide sequence(s), so as to; (a) reduce or minimise amplitude (ie peaks) in hydrophobicity across the length of the polypeptide, and/or (b) reduce or minimise the total length of any hydrophobic region(s) within the polypeptide.

Summary of the Invention:

In a first aspect, the present invention provides a method for designing a candidate polypeptide for expression in a bacterial host, said method comprising,

identifying one or more hydrophobic peptide sequences within a polypeptide of interest, and

arranging or re-locating at least one of said hydrophobic peptide sequences within said polypeptide so as to generate said candidate polypeptide with reduced or minimised amplitude in hydrophobicity and/or length of any hydrophobic region(s).

Preferably, the polypeptide of interest is non-native to the intended bacterial host. Since the most preferred bacterial host is *E. coli*, most preferably the polypeptide is non-native to *E. coli*.

The polypeptide of interest will preferably be a non-natural polypeptide or even a theoretical non-natural polypeptide (ie a polypeptide yet to be synthesised or expressed) comprising a plurality of amino acid sequences of interest some of which may be hydrophobic or suspected to be hydrophobic, and which has been found not to be, or is suspected not to be, efficiently expressed in a bacterial host. For such a polypeptide of interest, the method of the first aspect provides the possibility of identifying one or more hydrophobic peptide sequences, if any, within the polypeptide of interest and arranging or re-locating at least one of the hydrophobic peptide sequence(s) so as to generate a

candidate polypeptide with reduced or minimised amplitude in hydrophobicity and/or length of any hydrophobic region(s), and therefore an increased probability of being efficiently expressed in a bacterial host.

Preferably, the polypeptide of interest may be a synthesised or theoretical polyepitope polypeptide comprising a tandem array of epitopes of interest (eg CTL epitopes, which, as is mentioned above, often predominantly comprise hydrophobic amino acids). For such a polypeptide of interest, the method of the first aspect permits the design of candidate polyepitope polypeptides comprising a large number of epitopes of interest (eg 10 to 35 or more) with an increased probability of being efficiently expressed in a bacterial host, by enabling the possibility of identifying one or more hydrophobic epitopes and arranging or re-locating at least one of the hydrophobic epitope(s), so as to generate a candidate polyepitope polypeptide with reduced or minimised amplitude in hydrophobicity and/or length of any hydrophobic region(s).

It has been found that the method of the first aspect is best applied to the design of a candidate polyepitope polypeptide in a manner which identifies and ranks the relative hydrophobicity of each of the selected epitopes (nb The epitopes of interest may be a range of epitopes from a single pathogen (eg EBV) selected to provide a polyepitope polypeptide that covers the HLA diversity of the target population. The epitopes of interest may also be one or more epitopes from a range of pathogens or the epitopes may be derived from a non-microbial source such as a tumour cell for treating or preventing cancer.), groups the ranked epitopes into three groups of substantially equivalent numbers, based upon the identified relative hydrophobicity (ie so as to produce the groups, Group 1 = most hydrophobic, Group 2 = middle hydrophobicity, and Group 3 = least hydrophobic and "residual" epitopes where the total number of epitopes is not wholly divisible by 3), and then arranges the epitopes into triplets where the triplets contain an epitope from each group (ie three linked epitopes; epitope 1 - epitope 2 - epitope 3) and arranged into a candidate polyepitope polypeptide having the formula, Triplet 1 - Triplet 2 - .... - Triplet N, as follows:

	Epitope 1	Epitope 2	Epitope 3
Triplet 1 (N-terminal)	most hydrophilic of Group 2	most hydrophobic of Group 1	Most hydrophilic of Group 3
Triplet 2	2nd most hydrophilic of Group 2	2nd most hydrophobic of Group 1	2nd most hydrophilic of Group 3
Triplet N (C-terminal)	most hydrophobic of Group 2	most hydrophilic of Group 1	Most hydrophobic of Group 3

(Any "leftover" epitope(s) (ie least hydrophilic epitope(s) of Group 3) may be added to the C-terminal of Triplet N, or otherwise may be located within the candidate polyepitope polypeptide sequence so as to reduce any local peaks of hydrophobicity.)

5   Between the epitope triplets, or between any or all of the epitopes within a triplet, there may be intervening sequences (preferably short sequences of 1 to 10 amino acids) which may optionally be hydrophilic (eg lysine-lysine) so as to reduce any local peaks of hydrophobicity. Preferably, the epitopes within a triplet are contiguous.

10   Other simple methods for arranging the epitope(s) so as to minimise extremes in hydrophobicity in a polyepitope polypeptide will be readily apparent to persons skilled in the art, and are to be considered as forming part of the present invention (eg polyepitope polypeptide sequences based on sets of 4 epitopes could be constructed in a similar manner to that described above for triplets).

15   Once a candidate polypeptide has been designed in accordance with the method of the first aspect, a polynucleotide encoding the candidate polypeptide may be synthesised according to any of the methods well known to persons skilled in the art. The encoding polynucleotide may be incorporated into, for example, vectors such as viral vectors (eg vaccinia to provide a recombinant polyepitope viral vaccine) or expression vectors such as

20   those suitable for expression in a bacterial host.

Thus, in a second aspect, the present invention provides a method of expressing a polypeptide in a bacterial host, said method comprising,

designing a polypeptide in accordance with the method of the first aspect,

introducing a polynucleotide encoding said polypeptide into said bacterial host,

25   such that said bacterial host is capable of expressing said polypeptide, and

culturing said bacterial host under conditions suitable for expression of said polypeptide.

The expressed polypeptide may be isolated by, for example, lysing the bacterial host and purifying the polypeptide from the produced cell lysate.

5   The polynucleotide introduced into the bacterial host may encode the polypeptide in the form of a fusion of the polypeptide with a suitable carrier protein. Alternatively, the polypeptide could be expressed and subsequently linked to or otherwise associated with a suitable carrier protein. Suitable carrier proteins are well known to persons skilled in the art and include β-galactosidase, glutathione S-transferase and the gp350 structural protein

10   from EBV or a fragment thereof. The carrier protein may comprise additional useful epitopes. Further increases in expression benefits provided by ordering may be conferred by the carrier protein.

In a third aspect, the present invention provides a polypeptide designed in accordance with the method of the first aspect.

15   If desired, the polypeptide of the third aspect may be in the form of a fusion of the polypeptide with a suitable carrier protein.

In a fourth aspect, the present invention provides a polyepitope polypeptide designed in accordance with the method of the first aspect.

If desired, the polypeptide of the fourth aspect may be in the form of a fusion of the

20   polypeptide with a suitable carrier protein.

In a fifth aspect, the present invention provides a polyepitope polypeptide comprising N epitopes, wherein N is any integer preferably in the range of 5 to 50 (preferably, 10 to 35), said polyepitope polypeptide having the formula;

Triplet 1 - Triplet 2 - .... - Triplet N,

25   wherein each of said triplets comprises three linked epitopes selected by, .

identifying and ranking the relative hydrophobicity of each of the N epitopes,

grouping the ranked N epitopes into three groups of substantially equivalent numbers, based upon the identified relative hydrophobicity of the N epitopes, to produce a first group (ie Group 1) comprising the most hydrophobic epitopes, a second group (ie

30   Group 2) comprising the epitopes having a middle level of hydrophobicity, and a third group (ie Group 3) comprising the least hydrophobic epitopes, and

selecting the epitopes for each of said triplets according to the following table:

8

	Epitope 1	Epitope 2	Epitope 3
Triplet 1 (N-terminal)	most hydrophilic of Group 2	most hydrophobic of Group 1	Most hydrophilic of Group 3
Triplet 2	2nd most hydrophilic of Group 2	2nd most hydrophobic of Group 1	2nd most hydrophilic of Group 3
Triplet N (C-terminal)	most hydrophobic of Group 2	most hydrophilic of Group 1	Most hydrophobic of Group 3

Preferably, the first, second and third groups comprise identical numbers of epitopes. Where N is an integer not wholly divisible by 3 (ie an integer other than, for example, 6, 9, 12, 15, and 18), then the residual epitopes are preferably included within the
5   third group.

At the end of the step of selecting epitopes for each of said triplets, if there is/are any leftover epitope(s) (ie least hydrophilic epitope(s) of Group 3), then this/these may be added to the C-terminal of Triplet N, or otherwise may be located within the candidate polyepitope polypeptide sequence so as to reduce any local peaks of hydrophobicity.
10   Between the epitope triplets, or between any or all of the epitopes within a triplet, there may be intervening sequences (preferably short sequences of 1 to 10 amino acids) which may optionally be hydrophilic (eg lysine-lysine) so as to reduce any local peaks of hydrophobicity. Preferably, the epitopes within a triplet are contiguous.

If desired, the polyepitope polypeptide of the fifth aspect may be in the form of a
15   fusion of the polyepitope polypeptide with a suitable carrier protein.

In a sixth aspect, the present invention provides a polyepitope vaccine comprising a polyepitope polypeptide according to the fourth or fifth aspect and a pharmaceutically acceptable carrier and/or adjuvant.

In a seventh aspect, the present invention provides a polyepitope polypeptide
20   comprising an amino acid sequence substantially corresponding to:

FLRGRAYGL – PYLFWLAAI – HRCQAIRKK – RRIYDLIEL – VQPPQLTLQV-
GLCTLVAML - RLRAEAQVK -IEDPPFNSL –YLLEMLWRL – GQGGSPTAM -
AVLLHEESM - IALYLQQNWWTL-RAKFKQLL - SSCSSCPLSKI- TYGPVFMCL-
25   QAKWRLQTL- RPPIFIRRL- VSFIEFVGW - YPLHEQHGM –VEITPYKPTW -

9

CLGGLLTMV – EENLLDFVRF - TYSAGIVQI – LLDFVRFMGV -EGGVGWRHW

(SEQ ID NO:1);

or

FLRGRAYGL- PYLFWLAAI – HRCQAIRKK - RRIYDLIEL- GLCTLVAML–
RLRAEAQVK- IEDPPFNSL -TYSAGIVQI- LLDFVRFMGV- EGGVGWRHW-
IALYLQQNWWTL- RAKFKQLL - SSCSSCPLSKI- TYGPVFMCL- QAKWRLQTL-
RPPIFIRRL - VSFIEFVGW –YPLHEQHGM- VEITPYKPTW- CLGGLLTMV–
EENLLDFVRF - YLLEMLWRL – GQGGSPTAM - AVLLHEESM –VQPPQLTLQV

(SEQ ID NO:2).

In an eighth aspect, the present invention provides a polyepitope vaccine comprising a polyepitope polypeptide according to the seventh aspect and a pharmaceutically acceptable carrier and/or adjuvant.

Detailed Description of the Invention:

The present applicants have identified novel methods for designing a candidate polyepitope polypeptide, with an increased probability of being efficiently expressed in a bacterial host (ie in amounts detectable by SDS-PAGE). The method involves identifying one or more hydrophobic epitope(s) and arranging or re-locating at least one of the hydrophobic epitope(s) so as to generate a candidate polyepitope polypeptide with reduced or minimised amplitude in hydrophobicity and/or length of any hydrophobic region(s).

An algorithm to calculate hydrophobicity values of amino acid sequences and subsequently arrange sequences to; (a) reduce or minimise amplitude in hydrophobicity, and/or (b) reduce or minimise the length of hydrophobic sequences, was generated and applied to 26 CTL epitope sequences from EBV. This resulted in the design of two initial candidate polyepitope polypeptides (designated PT26A and PT26B, described hereinafter), one of which proved to be efficiently expressed in E. coli. The expressed polyepitope polypeptide shows promise as the basis of an EBV vaccine for prevention or treatment of infectious mononucleosis and/or EBV-related cancers such as Burkitts lymphoma, Hodgkin's disease, non-Hodgkin's lymphoma, naso-pharyngeal carcinoma, gastric adenocarcinoma, lymphomas associated with immunosupression, lymphoepithelioma-like carcinomas, and immunodeficiency-related leiomyosarcoma.

While looking for an explanation as to why the expression capabilities of *E. coli* for the two similar candidate polyepitope polypeptides were different, summations of hydrophobicity values (designated Hydrophobic Index (HI) values) were calculated for different numbers of epitopes over the length of the candidate polyepitope polypeptides to

5 identify local areas of hydrophobicity. Summation over 3 and 4 epitopes showed that there were regions in the non-expressed polypeptide where the HI value was higher than in the expressed polypeptide. This information enabled the identification of a threshold HI value, such that polypeptide sequences which comprised a region with an HI value in excess of the threshold value, could be predicted as being less likely of being efficiently

10 expressed in a bacterial host.

Thus, in a preferred embodiment of the methods of present invention for designing candidate polypeptides, the methods involve initially calculating hydrophobicity values and arranging peptide sequences to; (a) reduce or minimise amplitude in hydrophobicity, and (b) reduce or minimise the length of hydrophobic sequences; and then "fine-tuning", if

15 necessary, by calculating the HI values over different peptide sequence groups, thus providing numerical values for comparison and prediction of the likelihood of a candidate polypeptide sequence being efficiently expressed in a bacterial host. So, in applying this preferred embodiment to the design of a candidate polyepitope polypeptide, the method involves:

20 (i)   Calculating the hydrophobic value for each epitope using a suitable algorithm (eg Fauschere and Pliska, 1983 contained within the software package "Pinsoft 2" from Mimotopes Pty Ltd, Clayton, Victoria).

(ii)   Ranking the set of epitopes in order of decreasing hydrophobicity.

(iii)   Dividing the rank ordered set of epitopes into 3 equal groups (ie group 1 = most

25 hydrophobic, group 2 = middle hydrophobicity and group 3 = least hydrophobic (most hydrophilic)), and including any residual epitopes (ie epitopes left over after the set is divided by 3) in group 3 (ie the most hydrophilic group).

(iv)   Creating triplets of epitopes by taking the most hydrophilic of group 2 (ie last in group 2), then the most hydrophobic epitope (ie number 1 in group 1) and lastly the most

30 hydrophilic (ie last in group 3) until all epitopes in groups 1 and 2 have been used (nb "Leftover" epitopes are handled as set out in step (ix) below).

(v)   Arranging the triplets into a sequence in the order in which they were produced (ie Triplet 1 - Triplet 2 – Triplet 3 – etc).

(vi)   Plotting the hydrophobicity of the arranged triplet sequence using a suitable

algorithm (eg Fauschere and Pliska, 1983, or Hopp and Woods, 1981).

(vii)   If necessary, reducing hydrophobic amplitude by re-locating triplets from areas of low hydrophobicity into areas of high peak hydrophobicity and/or by re-locating individual hydrophobic (ie group 1) epitopes from areas of peak hydrophobicity into areas 5 of low hydrophobicity (nb The latter should not generally be necessary as a consequence of steps (iii) and (iv) above).

(viii)   Re-calculating the hydrophobicity plots and continuing to shuffle triplets as in step (vii) above to generate a final sequence arrangement.

(ix)   Placing any leftover epitopes (ie least hydrophilic epitopes of group 3) at the C-10 terminal of the final sequence arrangement or other location so as to further reduce local peaks in hydrophobicity (ie by inserting them adjacent to epitopes of peak hydrophobicity).

(x)   Placing any affinity tags (usually hydrophilic, eg a hexa-histidine sequence) at either the N- or C- terminal of the final polyepitope polypeptide sequence or at the C-15 terminal if the final polyepitope polypeptide sequence is to be expressed as a fusion protein.

The HI values may be calculated by using the mathematical expression:

$$e=m+n-1$$
$$HI_m = \sum_{e=m} x_e$$

20
(where m = group number, n = group size, e = integer (epitope) number, x = epitope hydophobicity value).

Preferably, the HI values are calculated using this mathematical expression when n=3 and n=4. In the examples provided hereinafter, this calculation predicted that to be 25 able to express linked, random, short amino acid sequences in *E. coli* in SDS-PAGE detectable amounts, the hydrophobic index over groups of three epitope sequences would need to be less than 1.8 ($HI_3 < 1.8$) and/or that over groups of four epitope sequences, the hydrophobic index would need to be less than 2.5 when x was calculated using Pinsoft 2 (Mimotopes Pty Ltd) and specifying the N-terminus as N-acetyl and the C-terminus as 30 carboxy amide. Different cut-off values will be obtained with different hydrophobicity algorithms.

It will be readily appreciated that the calculation of HI values in this manner, would be useful for predicting whether a natural, non-bacterial polypeptide or a derivative thereof may be efficiently expressed in *E. coli*. The present invention therefore further

12

extends to a method of predicting efficient expression of a polypeptide in a bacterial host, involving calculating HI values in accordance with the mathematical expression:

$$e = m+n-1$$

$$HI_m = \Sigma\, x_e$$

$$e = m$$

(where m = group number, n = group size, e = integer (epitope) number, x = epitope hydophobicity value).

The methods of the present invention permit the design of candidate polyepitope polypeptides comprising a large number of epitopes of interest (eg 10 to 35 or more) with an increased probability of being efficiently expressed in a bacterial host. The epitopes of interest may be a range of epitopes from a single pathogen selected to provide a polyepitope polypeptide that covers the HLA diversity of the target population, or the epitopes of interest may be one or more epitopes from a range of pathogens or tumour antigens. As is evident from the above, one particular application of the methods of the present invention is to the design of candidate polyepitope polypeptides comprising 26 EBV CTL epitopes for use in a vaccine to provide protection against EBV in over 90% of the human population. Another particular application of the methods of the present invention is to the design of candidate polyepitope polypeptides comprising CTL epitopes from cytomegalovirus (CMV), for use in a vaccine to prevent or treat CMV-causative diseases.

A candidate polypeptide designed in accordance with the methods of the present invention may be expressed by firstly synthesising a polynucleotide encoding the candidate polypeptide according to any of the methods well known to persons skilled in the art, and then by introducing the polynucleotide into a suitable bacterial host. Typically, this will be achieved by cloning the polynucleotide into an expression vector and then introducing the expression vector into a suitable bacterial host by any of the transformation methods well known to persons skilled in the art. Expression from the expression vector may result in the polypeptide being expressed as a fusion protein comprising the polypeptide and a suitable carrier protein (eg β-galactosidase, glutathione S-transferase or the gp350 structural protein from EBV or a fragment thereof). Alternatively, the polypeptide may be expressed by the bacterial host, and following isolation of the polypeptide, the polypeptide may be linked to or otherwise associated with a suitable carrier protein. The carrier protein may also confer additional useful properties (ie the carrier protein may comprise useful epitopes or sequences to enhance solubility, further enhance purification procedures or facilitate association with an adjuvant).

It is also contemplated that candidate polypeptides designed in accordance with the methods of the present invention may be readily expressed in whole cell lysates and non-bacterial host cells as well, and accordingly such alternative expression methods for candidate polypeptides are to be considered as forming part of the present invention. In

5 particular, the present invention is to be considered as extending to a method of expressing a polypeptide in a non-bacterial host cell such as a mammalian cell (eg a CHO cell or COS cell line), a yeast cell (eg *Saccharomyces cerevisiae*) or insect cell (eg SF9 cell line), wherein the method comprises designing a polypeptide in accordance with the method of the first aspect, introducing a polynucleotide encoding the polypeptide into the host cell such that

10 the host cell is capable of expressing the polypeptide, and culturing the host cell under conditions suitable for expression of the polypeptide. The expressed polypeptide may be isolated from the host cell culture by lysing the cells and purifying the polypeptide from the produced cell lysate, or alternatively, the polypeptide could be expressed with a suitable secretion signal such that the polypeptide is secreted into the culture medium

15 (from where it may be purified). Designing a polyepitope polypeptide in accordance with the methods of the present invention may also overcome non-secretion problems which are sometimes experienced when a hydrophobic polypeptide is expressed with a foreign secretion signal.

Where the expressed polypeptide is of pharmacological or veterinary significance,

20 the polypeptide may be formulated into a pharmaceutical or veterinary composition. Generally, such compositions will comprise a pharmaceutically acceptable or veterinary acceptable carrier, and may include other substances and excipients as may be required.

Polyepitope polypeptides may be formulated into vaccine compositions. Generally, such compositions will comprise a pharmaceutically acceptable or veterinary acceptable

25 carrier and may include adjuvants (eg an ISCOM™ adjuvant, DEAE, polysaccharides, liposomes and virus-like particles), and other substances and excipients as may be required. For example, the vaccine compositions may include immunomodulatory compounds (eg cytokines), and other proteins/compounds (eg melittin or regulatory proteins). The vaccine compositions may also include helper epitopes/CD4 epitopes or B-

30 cell epitopes. The vaccine compositions may be adapted for administration to a subject by, for example, intramuscular injection, nasal administration via an aerosol spray, or oral administration. Preferably, the vaccine compositions are ISCOM™ adjuvant compositions.

Polyepitope polypeptides may also be administered to a subject in the form of a viral vaccine (eg a recombinant polyepitope vaccinia or adenovirus) or DNA vaccine.

14

Thus, in a further aspect, the present invention provides a polynucleotide vaccine comprising a polynucleotide encoding a polypeptide designed in accordance with the method of the first aspect, and a pharmaceutically acceptable carrier and/or adjuvant.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The term "substantially corresponding" as used herein in relation to an amino acid sequence is intended to encompass the exact amino acid sequence as well as minor variations which do not result in a decrease in the biological activity of the amino acid sequence (eg variations which do not diminish the ability of an epitope to provoke a CTL immune response). These variations may include one or more conservative amino acid substitutions. The conservative amino acid substitutions envisaged are: G, A, V, I, L, M; D, E, N, Q; S, C, T; K, R, H; and P, Nα-alkylamino acids.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is therefore, not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the art relevant to the present invention as it existed in Australia or elsewhere before the filing or priority date of the present specification.

The invention will hereinafter be further described by way of the following non-limiting examples and accompanying figures.

Brief Description of the accompanying Figures:

Figure 1 provides the epitope configuration and amino acid sequences for EBV-polyepitope polypeptides, PT26A and PT26B. Numbers above epitopes represent hydrophobicity value for each epitope calculated with Pinsoft 2 software, specifying an acetyl N-teminal and an amide C-terminal.

Figure 2 provides hydrophobicity plots of PT26A and PT26B. Hydrophobicity values of a moving nine amino acid window are derived using the algorithm of Fauchere and Pliska, 1983.

Figure 3 provides the amino acid sequence of the fusion of residues 21 to 447 of EBV gp350 to PT26A (A) or to PT26B (B).

Figure 4 shows Coomassie stained SDS-PAGE gels showing the time course of expression following induction with IPTG (+IPTG or I) of: (A) PT26A (at approximately 30kDa), (B) gp350/PT26A (at approximately 80kDa), and (C) PT26B expression. Arrows indicate the location of recombinant protein.

Figure 5 provides ELISPOT assay results of CTL responses to the five HLA A2 epitopes contained within the polyepitope polypeptides, PT26A and PT26B, under two formulation conditions: (A) 10µg gp350-PT26A, (B) 10µg gp350-PT26B, and (C) positive control peptide mix containing a mixture of each of the 5 A2 epitopes contained in the polypeptide polypeptide. Each of the epitopes is represented below by their first 3 amino acids. The CTL response of each mouse M1-M5 to each A2 epitope is presented as a bar indicating the number of IFN-γ spots produced.

Example 1: EBV polyepitope fusions as vaccine candidates.

MATERIALS AND METHODS

Epitope sequences

The 26 CTL epitopes for inclusion in an EBV vaccine, the proteins from which they originate and HLA type are shown in Table 1.

TABLE 1. CTL epitopes included in the EBV polytopes

HLA type	EBV protein	Epitope
A2	LMP2	CLGGLLTMV
	BMLF1	GLCTLVAML
	EBNA6	LLDFVRFMGV
	LMP1	YLLEMLWRL
	LMP1	YLQQNWWTL
A3	EBNA3	RLRAEAQVK
A11	LMP2	SSCSSCPLSKI
A23	LMP2	PYLFWLAAI
A24	LMP2A	TYGPVFMCL
	EBNA4	TYSAGIVQI
B7	EBNA3	RPPIFIRRL
B8	EBNA3	FLRGRAYGL
	EBNA3	QAKWRLQTL
	BZLF1	RAKFKQLL
B27	EBNA4	HRCQAIRKK
	EBNA6	RRIYDLIEL
B35	EBNA4	AVLLHEESM
	EBNA3	YPLHEQHGM
B44	EBNA6	EENLLDFVRF
	EBNA6	EGGVGWRHW
	EBNA4	VEITPYKPTW
B46	EBNA3	VQPPQLTLQV
B57	LMP1	IALYLQQNW
B58	EBNA4	VSFIEFVGW
B60	LMP2	IEDPPFNSL
B62	EBNA4	GQGGSPTAM

Design/Ordering of epitopes

The following method was used to generate ordered arrangements of CTL epitopes to produce a polyepitope sequence with favourable hydrophobicity characteristics:

5   (i)   The hydrophobic value for each epitope was calculated using a suitable algorithm (ie Pinsoft 2 from Mimotopes Pty Ltd, specifying the N-terminus as N-acetyl and the C-terminus as carboxy amide).

(ii)   The set of epitopes was ranked in order of decreasing hydrophobicity.

10

(iii)   The rank ordered set of epitopes was divided into 3 equal groups (ie group 1 = most hydrophobic, group 2 = middle hydrophobicity and group 3 = least hydrophobic (most hydrophilic)). Residual epitopes (ie 2 epitopes left over after the set of 26 was divided by 3), were included in the most hydrophilic group.

15

(iv)   Triplets of epitopes were created by taking the most hydrophilic of group 2 (ie last in group 2), then the most hydrophobic epitope (ie number 1 in group 1) and lastly the most hydrophilic (ie last in group 3). This was continued until all epitopes in groups 1 and 2 had been used (nb "Leftover" epitopes were added to the C-terminal end of the final sequence arrangement).

20   sequence arrangement).

(v) .   The triplets were then arranged into a sequence in the order in which they were produced (ie Triplet 1 – Triplet 2 – Triplet 3 – etc .

25   (vi)   The hydrophobicity of this triplet arrangement was then plotted using a suitable algorithm (ie Fauschere and Pliska)).

(vii)   If and where necessary, relocating triplets from areas of low hydrophobicity into areas of high peak hydrophobicity in order to reduce hydrophobic amplitude.

30

(viii)   Re-calculating the hydrophobicity plots and continuing to shuffle triplets as in the step (vii) above. (nb It is possible to move individual hydrophobic (ie group 1) epitopes from areas of peak hydrophobicity into areas of low hydrophobicity or vice-versa, to further reduce peak hydrophobicity amplitude, however, this should not be necessary as

the ordering process above minimises peaks).

(ix)   Any residual epitopes (ie least hydrophilic of group 3) can be placed C-terminally in the final sequence arrangement or can be used to further reduce local peaks in hydrophobicity by inserting them adjacent to epitopes of peak hydrophobicity, according to a hydrophobicity plot of the assembled triplets.

(x)   Any affinity tags (usually hydrophilic, eg a hexa-histidine sequence) should be located either N- or C- terminally or preferably C-terminally if the construct is a fusion protein.

(xi)   Confirmation of satisfactory HI.

The example of PT26A is shown in Table 4.

Using this process, EBV polyepitope configurations PT26A and PT26B were created.

Hydrophobic Index (HI) calculations

HI values for favourable configurations (PT26A, PT26B) were calculated according to the mathematical expression:-

$$e=m+n-1$$
$$HI_m = \Sigma\, x_e$$
$$e=m$$

(where m = group number, n = group size, e = integer (epitope) number, x = epitope hydophobicity value) over values of n from 2 to 5.

Preparation of recombinant proteins

The DNA sequences encoding the polyepitopes (PT26A and PT26B) were generated from synthetic oligonucleotides using a Splicing by Overlap Extension technique (SOE) as described by Horton et al (1995). The codon usage was optimised for *E. coli* expression (Wada et al 1992). The polyepitopes were tagged at the C-terminus with a hexa-histidine tag for protein purification and detection. The DNA was subcloned into pET28b (Novagen) and transformed into *E. coli* BL21(DE3) cells (Novagen) for expression.

19

A fragment corresponding to the N-terminal region (amino acid residues 21 – 447) of EBV gp350 was amplified from plasmid DNA containing the full length gp350 sequence by PCR using the following oligonucleotides:

5' AGGGATCCCATGGAAGATCCTGGTTTTTTC 3' (forward) and

5   5' TCTAGAGGTCGACACCTGTCGTTGTATTGGG 3' (reverse). This DNA fragment was subcloned into pET28b (Novagen) containing the polyepitope insert, resulting in an in-frame fusion between gp350 and the polyepitope polypeptide. The constructs were referred to as gp350/PT26A and gp350/PT26B.

For protein expression testing, transformed cells were grown in 50mg/ml

10   Kanamycin containing L broth at 37°C until OD600 reached 2. Protein expression was induced by the addition of IPTG (0.25mM) and cells were grown for another 3hr. Cells were pelleted and boiled in SDS sample buffer before analysing by SDS-PAGE.

For protein purification, cells induced by IPTG were pelleted, resuspended in binding buffer (20mM Tris-HCl pH 7.9, 0.5M NaCl, 5mM imidazole) and then sonicated.

15   Inclusion bodies were pelleted and washed in buffer. The proteins were solubilised overnight in binding buffer containing 8M urea and purified on a Ni^{++}-NTA column.

Preparation of ISCOM™ formulations

ISCOMATRIX™-adjuvant was prepared by combining adjuvant components in a

20   formulation vessel. Cholesterol, 1,2, dipalmitoyl phosphocholine (DPPC), and ISCOPREP™ as source of purified Quillaja saponins, were mixed in a ratio of 1:1:5 in the presence of the detergent Mega-10 (United States Patent No. 5,679,354) at a concentration of 2%. The detergent was removed by diafiltration with PBS and the formation of ISCOMATRIX™ confirmed by negative contrast electron microscopy revealing complexes.

25   including cage like structures with a diameter of 40nm. ISCOM™-adjuvanted vaccines were prepared by mixing the EBV polyepitope antigen with preformed ISCOMATRIX™-adjuvant, which was prepared as described below: The dose strength of ISCOM-adjuvant as saponin was quantified by reverse phase HPLC assay.

ISCOM™ vaccines were prepared by gentle mixing at 22°C of an equal volume of

30   2x final dose strength ISCOMATRIX™ with an equal volume of 2x final dose strength EBV polyepitope antigen (gp350-PT26A and B). After 60 minutes, the formulation was subjected to extensive dialysis, in order to remove urea, at 4°C into PBS buffer pH6.2 using

12,000 molecular weight cut off dialysis membrane (Cellu Sep T3, San Antonio TX).

Mouse immunogenicity

Dosing

5   Female HLA A2 transgenic C57Bl/6 mice (HDD) were bred at Queensland Institute of Medical Research (QIMR) and immunised at 5-7 weeks of age. Mice were housed in filter-topped cages in the PC3 animal facility at QIMR. Groups of 4 or 5 mice were dosed sub-cutaneously at the tail base with 0.1ml formulation. This was followed by removal of spleens at day 21 for *ex-vivo* ELISPOT assay (below).

10   Mice dosed sub-cutaneously, received 10µg ISCOM™-adjuvant (as saponin) and 10µg EBV polyepitope antigen. For a vaccine control group, mice were dosed sub-cutaneously with a peptide mixture comprising free peptides (Mimotopes Pty Ltd) formulated with tetanus toxoid and Montanide ISA 720 (SEPPIC, Paris, France) as previously described (Elliot et al, 1999). Peptide control immunisations come from two

15   groups of mice, one group immunised with GLCTLVAML/YLLEMLWRL/LLDFVRFMGV peptide mixture and the other YLQ/CLG. A mixture of all five epitopes showed some insolubility problems.

CTL activity (ex-vivo ELISPOT)

20   *Ex vivo* ELISPOT measures/quantitates both effector and memory CTL which secrete IFNγ. Peptide-specific IFNγ secreting cells are enumerated by an enzyme linked immuno-spot (ELISpot) assay modified from Murali-Krishna *et al*. Flat bottomed 96-well microtitre plates are coated overnight with 5 ug/mL of rat anti-mouse IFNγ antibody(clone RA-6A2, BD PharMingen, San Diego, California). Coated plates are then blocked for 1

25   hour with 1% foetal FBS in PBS, and then washed three times with PBS/0.05% Tween 20 (PBS-T), and incubated for 1 h at 37°C with medium comprising RPMI 1640, supplemented with 100 µg/ml streptomycin and 100 IU/ml penicillin, 10% FBS and 10^{-5} M 2-mercaptoethanol. Mouse splenocytes were then treated with red blood cell lysis buffer, washed and resuspended to 1×10^7 cells/mL in medium, for use in the *ex vivo* IFNγ

30   ELISPOT assay. Splenocytes (1×10^6/well) are then placed in the first wells of the ELISpot plate and serially diluted two fold. Recombinant human IL-2 (kindly provided by Cetus Corp.,Emeryville, California) is then added to the plate at a final concentration of 5 IU/well together with EBV peptide (Mimotopes Pty Ltd) at a final concentration of 100

μg/ml. Media containing IL-2 without peptide is added to control wells. The final volume in each well is 100 μl. Plates are incubated at 37°C in 5% CO_2 for approximately 18 hours. After incubation, cells are lysed by rinsing the plates in H_2O and then washed twice in PBS-T. Biotinylated anti-mouse IFNγ antibody (commercial source) is diluted 1:500 (2 μg/ml

5   final concentration) in PBS-T/5% FBS and added to all wells at 50 μl/well and incubated for 2 hours at RT. Plates are then washed in PBS-T and streptavidin-alkaline phosphatase, diluted 1:400 in PBS-T/5% FBS, is added at 50 μl per well and incubated for a further 2 hours. After washing, plates are developed by adding Sigma Fast BCIP/NBT substrate at 50 ul/well. Plates are incubated at 37°C for approximately 20 minutes to allow colour

10   development, and then washed with water to stop the reaction. IFNγ specific spots are counted using KS ELISPOT Reader (Zeiss).

RESULTS

Epitope Fusions

15   26 EBV CTL epitopes were selected to provide >90% human population coverage for a vaccine formulation.

In order to link these CTL epitopes (Table 1) together and facilitate the design of a polyepitope antigen to form the basis of an EBV prophylactic vaccine, the hydrophobicity values of each epitope were calculated using Pinsoft 2 software. Two versions of 26

20   epitopes were then ordered into configurations PT26A and PT26B (Figure 1), which minimised peak hydrophobicity and hydrophobic sequence length (Figure 2). When these constructs were cloned for expression in *E. coli* it was found that only one of the configurations (PT26A) was able to produce a polyepitope polypeptide (Figure 4A). PT26B was not produced (Figure 4C).

25   To identify a potential reason for this unexpected finding, local areas of high hydrophobicity were examined by summation of overlapping hydrophobicity values (Pinsoft 2, Mimotopes Pty Ltd) to provide a hydrophobicity index (HI) over varying numbers of peptides in a group (n). For n=2, no correlation was apparent. However, for n=3 and n=4 (Table 2), the highest HI values for the expressed sequence (PT26A) were

30   lower (1.79 and 2.51 respectively) than the highest values obtained for the non-expressed sequence (2.02 and 2.54 respectively). For n=5, again no significant differences were seen. This would indicate that there were local areas of slightly higher hydrophobicity in PT26B than PT26A.

The above analysis can be represented by the mathematical expression:-

$$HI_m = \sum_{e=m}^{e=m+n-1} x_e$$

5   (where m = group number, n = group size, e = integer (epitope) number, x = epitope hydophobicity value).

Overall, the results with n=3 and n=4 appeared to show the most distinctive differences. This would make the prediction that to be able to express linked, random, short amino acid sequences in *E. coli* in SDS PAGE detectable amounts, the hydrophobic

10   index over groups of three epitope sequences would need to be less than 1.8 ($HI_3 < 1.8$) and/or that over groups of four < 2.5, where x was calculated using Pinsoft 2 (Mimotopes Pty Ltd) and specifying the N-terminus as N-acetyl and the C-terminus as carboxy amide. Different cut-off values will be obtained with different hydrophobicity algorithms.

Table 3 shows the HI values over 3 peptides (n=3) for 15 random arrangements of

15   the 26 EBV CTL epitopes that were generated and analysed by calculating the HI. This shows that all 15 random configurations contain multiple HI values which are all in the range predicted to preclude the production of a recombinant polypeptide in *E. coli* (ie HI greater than 1.8).

TABLE 2:

3mers   (n=3)

PT26A	1.33	1.34	0.87	1.79	1.34	1.19	1.19	1.43	1.46	1.44	1.46	1.48	1.48	1.6	1.62	1.62	1.64	1.69	1.68	1.69	1.7	1.59	1.6
PT26B	1.33	1.34	1.17	1.18	1.19	0.87	1.64	1.6	1.9	1.39	1.48	1.48	1.6	1.62	1.62	1.64	1.69	1.68	1.69	2.02	1.38	1.46	1.16

4mers   (n=4)

PT26A	1.72	1.89	1.72	1.73	1.74	2.04	1.37	1.86	2.29	1.64	1.91	2.31	1.8	2.07	2.45	1.96	2.16	2.51	2.03	2.22	2.41	1.95
PT26B	1.72	2.19	1.11	1.58	1.72	1.58	2	2.43	2.1	1.84	2.31	1.8	2.07	2.45	1.96	2.16	2.51	2.54	2.2	1.81	2.01	

TABLE 3:

Hydrophobic Index

3mers (n=3)

																								# >1.8
PT26B	1.33	1.34	1.17	1.18	1.19	0.87	1.64	1.6	1.9	1.39	1.48	1.48	1.6	1.62	1.62	1.64	1.69	1.68	1.69	2.02	1.38	1.46	1.16	2
PT26A	1.33	1.34	0.87	1.79	1.34	1.19	1.19	1.43	1.46	1.44	1.46	1.48	1.48	1.6	1.62	1.62	1.64	1.69	1.68	1.69	1.7	1.59	1.6	0
Random																								
1	1.93	1.62	1.41	0.93	1	1.67	1.66	1.64	0.72	0.91	0.82	1.21	1.51	1.09	1.29	0.89	1.5	2	2.42	2.25	2.05	2.03	1.85	7
2	0.69	0.64	1.54	1.65	1.75	1.31	1.39	0.88	1.51	1.81	2.58	2.41	1.94	1.75	1.24	1.69	1.49	1.98	1.58	1.44	0.81	0.91	1.55	5
3	1.67	1.47	1.62	2	1.52	1.1	0.45	0.96	1.58	1.8	1.33	1.61	1.74	1.88	1.9	1.8	1.67	0.76	0.8	1.24	1.62	1.86	1.39	5
4	1.86	1.5	1.71	1.28	1.36	1.19	1.17	1.56	2.05	2.52	2.15	2.13	1.21	1.78	1.15	1.69	0.85	0.59	0.51	1.16	1.56	1.69	1.41	5
5	1.67	1.36	1.44	1.88	1.68	1.58	0.93	0.99	0.49	0.51	0.98	1.88	2.08	2.25	2.13	2.2	1.53	1.11	1.61	1.61	2.03	1.55	1.25	7
6	1.1	0.94	1.39	1.46	2.11	1.46	1.58	1.22	1.47	1.74	1.62	2.02	1.91	2.58	2.05	1.87	1.23	1.73	1.28	1.28	1.28	1.15	1.15	6
7	1.03	0.58	0.72	1.34	1.6	1.22	0.85	1.05	0.94	1.14	1.06	1.22	1.27	1.78	2.11	2.02	2.21	2.2	2.18	1.63	1.81	1.8	2.18	7
8	1.48	1.88	1.52	2.03	1.7	1.41	1.25	0.66	0.82	0.65	1.74	1.75	1.31	0.68	1.16	1.65	1.79	1.78	2.18	1.97	1.51	1.53	1.61	4
9	2.04	1.83	1.75	1.16	1.02	1.45	1.85	2	2.01	1.56	1.74	1.74	1.76	1.17	1.19	1.15	1.53	1.39	0.96	0.82	0.56	1.48	1.82	6
10	1.24	1.15	0.71	1.62	1.69	2.13	1.68	1.57	1.29	1.24	1.72	2.02	1.85	1.74	1.42	2.26	1.91	1.81	1.11	1.58	1.35	1.55	0.65	6
11	0.52	0.94	0.86	1.03	1.1	1.23	1.26	1.66	2.06	1.88	1.52	1.5	1.82	2.18	2.37	2.38	1.48	0.85	1.17	1.95	1.9	1.6	1.71	8
12	1.56	1.86	1.97	1.46	1.4	2.05	1.4	1.16	0.72	0.26	0.88	1.96	1.92	1.77	1.58	2.06	2.02	1.53	0.88	0.99	1.5	1.84	1.82	9
13	0.68	0.74	1.36	1.83	1.72	0.92	1.57	1.4	1.35	1.7	1.35	1.32	1.64	2	2.51	2.39	2.37	2.08	1.31	1.48	1.16	1.6	1.01	6
14	1.4	1.35	1.36	1.17	1.62	2.06	1.62	1.24	1.39	1.2	0.91	0.7	0.84	1.18	1.41	1.92	2.23	2.51	2.09	1.62	1.14	1.53	1.88	6
15	0.52	1.03	1.67	2	2.02	1.85	1.67	1.63	1.61	1.66	1.55	1.27	1.7	1.77	1.74	1.25	0.79	1.09	1.58	2.5	1.86	1.74	1.25	5

TABLE 4: Ordering process for generation of a polyepitope polypeptide for 26* EBV CTL epitopes (PT26A).

The hydrophobicity value for each epitope is calculated, then the epitopes are rank ordered, and grouped into triplets. The sequence is assessed using a hydrophobicity plot and scored for HI value by summing the epitope hydrophobicity values over a moving 3-mer window. If necessary, fine-tuning of the epitope order is done and the sequence reassessed. The final epitope order and amino acid sequence for the 26* EBV CTL epitopes and hexa-histidine affinity tag is shown below.

HLA Type	EPITOPE	Hydrophobicity (Pinsoft 2)	Rank ordered on hydrophobicity	Hyd.	Grouped into triplets	Hyd.	Order after fine tuning	Hyd.	Sum hydrophobicity for triplets
A2	CLGGLLTMV	0.82	PYLFWLAAI	1.02	FLRGRAYGL	0.38	FLRGRAYGL	0.38	
A2	GLCTLVAML	0.85	GLCTLVAML	0.85	PYLFWLAAI	1.02	PYLFWLAAI	1.02	
A2	LLDFVRFMGV	0.71	YLLEMLWRL	0.85	HRCQAIRKK	-0.07	HRCQAIRKK	-0.07	1.33
A2	YLLEMLWRL	0.85	IALYLQQNWWTL*	0.83	RRIYDLIEL	0.39	RRIYDLIEL	0.39	1.34
A2/B57	IALYLQQNWWTL*	0.83	TYGPVFMCL	0.83	GLCTLVAML	0.85	VQPPQLTLQV	0.55	0.87
A3	RLRAEAQVK	-0.06	VSFIEFVGW	0.83	RLRAEAQVK	-0.06	GLCTLVAML	0.85	1.79

Allele	Score	Peptide	Score	Peptide	Score	Peptide	Score	Peptide	
A1	0.45	SSCSSCPLSKI	0.82	CLGGLLTMV	0.40	IEDPPFNSL	-0.06	RLRAEAQVK	1.34
A23	1.02	PYLFWLAAI	0.71	LLDFVRFMGV	0.85	YLLEMLWRL	0.40	IEDPPFNSL	1.19
A24	0.53	TYSAGIVQI	0.55	VQPPQLTLQV	0.18	GOGGSPTAM	0.85	YLLEMLWRL	1.19
A24	0.83	TYGPVFMCL	0.53	TYSAGIVQI	0.43	AVLLHEESM	0.18	GOGGSPTAM	1.43
B7	0.47	RPPIFIRRL	0.52	VEITPYKPTW	0.83	IALYLQQNWWTL*	0.43	AVLLHEESM	1.46
B8	0.38	FLRGRAYGL	0.47	RPPIFIRRL	0.20	RAKFKQLL	0.83	IALYLQQNWWTL*	1.44
B8	0.32	QAKWRLQTL	0.45	SSCSSCPLSKI	0.45	SSCSSCPLSKI	0.20	RAKFKQLL	1.46
B8	0.20	RAKFKQLL	0.43	AVLLHEESM	0.83	TYGPVFMCL	0.45	SSCSSCPLSKI	1.48
B27	-0.07	HRCQAIRKK	0.40	IEDPPFNSL	0.32	QAKWRLQTL	0.83	TYGPVFMCL	1.48
B27	0.39	RRIYDLIEL	0.39	RRIYDLIEL	0.47	RPPIFIRRL	0.32	QAKWRLQTL	1.60
B35	0.34	YPLHEQHGM	0.38	FLRGRAYGL	0.83	VSFIEFVGW	0.47	RPPIFIRRL	1.62
B35	0.43	AVLLHEESM	0.36	EGGVGWRHW	0.34	YPLHEQHGM	0.83	VSFIEFVGW	1.62
B44	0.52	VEITPYKPTW	0.35	EENLLDFVRF	0.52	VEITPYKPTW	0.34	YPLHEQHGM	1.64
B44	0.36	EGGVGWRHW	0.34	YPLHEQHGM	0.82	CLGGLLTMV	0.52	VEITPYKPTW	1.69

B44	EENLLDFVRF	0.35	QAKWRLQTL	0.32	EENLLDFVRF	0.35	CLGGLLTMV	0.82	1.68
B46	VQPPQLTLQV	0.55	RAKFKQLL	0.20	TYSAGIVQI	0.53	EENLLDFVRF	0.35	1.69
B58	VSFIEFVGW	0.83	GQGGSPTAM	0.18	LLDFVRFMGV	0.71	TYSAGIVQI	0.53	1.70
B60	IEDPPFNSL	0.40	RLRAEAQVK	-0.06	EGGVGWRHW	0.36	LLDFVRFMGV	0.71	1.59
B62	GQGGSPTAM	0.18	HRCQAIRKK	-0.07	VQPPQLTLQV	0.55	EGGVGWRHW	0.36	1.60
	HHHHHH	0.04	HHHHHH	0.04	HHHHHH	0.04	HHHHHH	0.04	1.11

(* The epitope IALYLQQNWWTL is comprised of two overlapping CTL epitopes IALYLQQNW and YLQQNWWTL that were combined for this study.)

28

Fusions of these polyepitopes with the N-terminal 400 amino acids of a naturally-occurring EBV protein (gp350).

Provision of CD4 help has previously been shown to improve CTL induction (Thuy et al, 2001) and the EBV structural protein gp350 was identified as the preferred candidate
5   to provide this property because it would provide cognate help. Hence the two polyepitopes, PT26A and PT26B, the latter of which was unable to be expressed in *E. coli*, were cloned onto the C-terminus of the N-terminal 400 amino acids of gp350 (Figure 3) and expressed as the fusion protein in *E. coli*. Both of these fused polypeptides were well-produced, being clearly visible on a Coomassie-stained gel above the background of *E. coli*
10   proteins (profile for PT26A shown in Figure 4B).

CTL activity

Vaccination of HLA A2 transgenic mice with either gp350-PT26A-ISCOM™ vaccine or gp350-PT26B-ISCOM™ vaccine induced a CTL immune response (IFN ELISPOT) to all
15   five A2 epitopes (Figure 5), thus indicating that all A2 epitopes were properly processed and presented to the immune system.

29

References:

Elliot S.L., Pye S., Le T. et al. Peptide based cytotoxic T-cell vaccines; Delivery of multiple epitopes, help, memory and problems. Vaccine 1999; 17:2009-2019.

5   Fauchere J.-L and Pliska V.E. (1983), Eur. J. Med. Chem. 18:369-375.

Fisher M-T, Yuan X. The rates of commitment to renaturation of rhodanese and glutamine synthetase in the presence of the groE chaperonins. J. Bio. Chem. 1994 Nov 25; 269(47): 29598-601.

10

Frace A-M, Klimov A-I, rowe T, Black R-A, Katz J-M. Modified M2 proteins produce heterotypic immunity against influenza A virus. Vaccine. 1999 May; 17(18): 2237-44.

Hayer-Hartl M-K, Ewbank J-J, Creighton T-E, Hartl F-U. Conformational specificity of
15   the chaperonin GroEL for the compact folding intermediates of alpha-lactalbumin. EMBO J. 1994 Jul; 13(13): 3192-202.

Hendrick J-P, Hartl F-U. The role of molecular chaperones in protein folding. FASEB J. 1995 Dec; 9(15): 1559-69.

20

Hobman T-C, Seto N-O, Gillam S. Expression of soluble forms of rubella virus glycoproteins in mammalian cells. Virus-Res. 1994 Mar; 31(3): 277-89.

Hopp T.P., Woods K.R. Proc. Nat. Acad. Sci USA 1981; 78:3824-3828.

25

Horton R-M, Cai Z-L, Ho S-N, Pease L-R. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques, 1990; 528-35.

Lin Z, Schwartz F-P, Eisenstein E. The hydrophobic nature of GroEL substrate
30   binding. J. Biol. Chem. 1995 ; 270(3): 1011-4.

Murali-Krishna K., Altman J.D., Suresh M. et al. Counting antigen specific CD8 T-cells; A re-evaluation of bystander activation during viral activation. Immunity 1998; 8:177-187.

5   Polte T, Newman W, Raghunathan G, Gopal TV. Structural and functional studies of full-length vascular cell adhesion molecule – 1: internal duplication and homology to several adhesion proteins. DNA Cell Biol. 1991; 10(5): 349-57

     Richarme G, Kohiyama M. Amino acid specificity of the Escherichia coli chaperone
10   GroEL (heat shock protein 60). J. Biol. Chem. 1994 Mar 11; 269(10): 7095-8.

     Thuy TT Le et al. Cytotoxic T cell polyepitope vaccines delivered by ISCOMS. Vaccine 2001 19:4669-4675.

15   United States Patent No. 5,679,354; Morein et al; Saponin preparations and use thereof in iscoms, 5 March 5 2002.

     Wada K, Wada Y, Ishibashi F, Gojobori T, Ikemura T. Codon usage tabulated from the Gen Bank genetic sequence data. Nucleic Acids Research 1992 20 Suppl:2111-8
20

     Zahn R, Pluckthun A. Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. II. GroEL recognizes thermally unfolded mature beta-lactamase. J. Mol. Biol. 1994 Sep 16; 242(2): 165-74.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

5

Dated this 12th day of July 2002

The Council of the Queensland Institute of Medical Research
By their Patent Attorneys:
Blake Dawson Waldron Patent Services

Figure 1

P26A
```
        0.38      1.02     -0.07      0.39      0.55      0.85     -0.06      0.4       0.85
    MVDFLRGRAYGL- PYLFWLAAI-HRCQAIRKK - RRIYDLIEL-VQPPQLTLQV-GLCTLVAML- RLRAEAQVK-IEDPPFNSL-YLLEMLWRL-
        0.18      0.43     0.83      0.2       0.45      0.83      0.32      0.47      0.83
    GQGGSPTAM - AVLLHEESM - IALYLQQNWWTL-RAKFKQLL - SSCSSCPLSKI- TYGPVFMCL-QAKWRLQTL- RPPIFIRRL- VSFIEFVGW-
        0.34      0.52     0.82      0.35      0.53      0.71      0.36
    YPLHEQHGM-  VEITPYKPTW-CLGGLLTMV-EENLLDFVRF-TYSAGIVQI-LLDFVRFMGV-EGGVGWRHW-HHHHHH
```

P26B
```
        0.38      1.02     -0.07      0.39      0.85     -0.06      0.4       0.53      0.71
    MVDFLRGRAYGL- PYLFWLAAI-HRCQAIRKK - RRIYDLIEL-GLCTLVAML-RLRAEAQVK- IEDPPFNSL-TYSAGIVQI- LLDFVRFMGV-
        0.36      0.83     0.2       0.45      0.83      0.32      0.47      0.83      0.34
    EGGVGWRHW - IALYLQQNWWTL- RAKFKQLL - SSCSSCPLSKI - TYGPVFMCL- QAKWRLQTL- RPPIFIRRL - VSFIEFVGW -YPLHEQHGM-
        0.52      0.82     0.35      0.85      0.18      0.43      0.55
    VEITPYKPTW-CLGGLLTMV-EENLLDFVRF - YLLEMLWRL - GQGGSPTAM - AVLLHEESM -VQPPQLTLQV-HHHHHH
```

Figure 2

Hydrophobicity Index (Fauschere & Pliska)

Position of 9-mer moving window
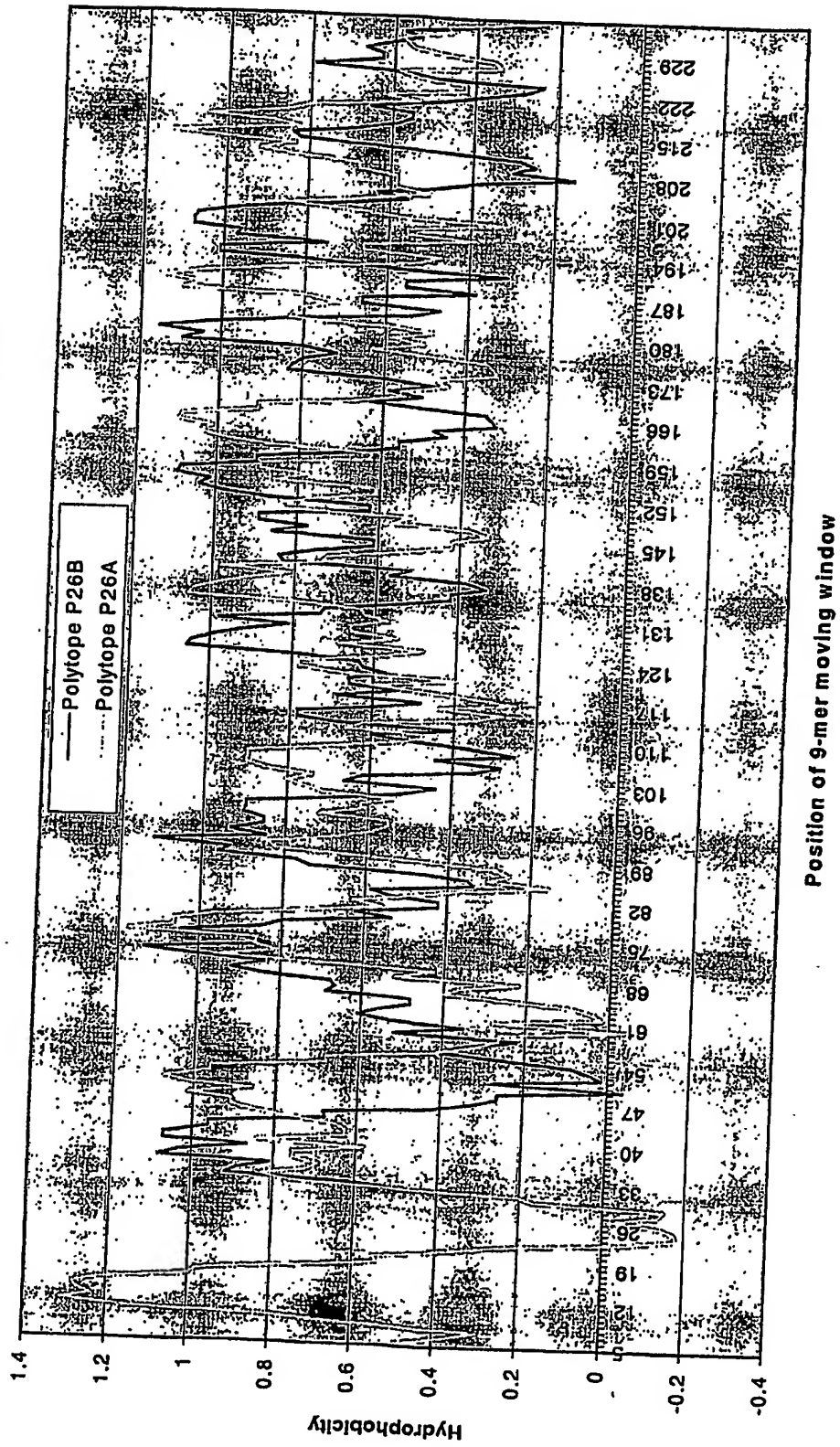
Figure 3A

MEDPGFFNVEIPEFPFYPTCNVCTADVNVTINFDVGGKKHQLDLDF<u>GQLTPHTKAVYQPRGAFGGSENAT</u>
<div align="right">T-help epitope   70</div>

NLFLLELLGAGELALTMRSKKLPINVTTGEEQQVSLESVDVYFQDVFGTMWCHHAEMQNPVYLIPETVPY
<div align="right">140</div>

IKW<u>DNCNSTNITAVVRAQGLD</u>VTLPLSLPTSAQDSNFSVKTEMLGNEIDIECIMEDGEISQVLPGDNKFN
    T-help epitope

ITCSGYESHVPSGGILTSTSPVATPIPGTGYAYSLRLTPRPVSRFLGNNSILYVFYSGNGPKASGGDYCI
<div align="right">210</div>

QSNIVFSDEIPASQDMPTNTTDITYVGDNATYSVPMVTSEDANSPNVTVTAFWAWPNNTETDFKCKWTLT
<div align="right">280</div>

SGT<u>PSGCENI</u>SGAFASNRTFDITVSGLGTAPKTLIITRTATNATTTTHKVIFSKAPESTTTSPTLNTTGF
<div align="right">350</div>

ADPNTTTG*VDF*<u>LRGRAYGL</u>PYLFW<u>LAAI</u><u>HRCQAIRKKRRIYDLIEL</u><u>VQPPQLTLQV</u><u>GLCTLVAML</u><u>RLRAE</u>
<div align="right">420</div>

  428   B8   A23   B27   B27   B46   A2   A3

<u>AQVKIEDPPF</u>NS<u>LYLLEMLWRL</u>GQGG<u>SPTAMAVLLHEESM</u>IA<u>LYLQQNWWT</u>L<u>RAKFKQLL</u>SSCSSCPLSK
  B60   A2   B62   B35   B57/A2   B8   A11

<u>ITYGPVFMCL</u>QAKW<u>RLQTL</u><u>RPPIFIRRL</u><u>VSFIEFV</u>GWY<u>PLHEQHGM</u>VEITPYKPTW<u>CLGGLLTMV</u>EENLL
  A24   B8   B7   B58   B35   B44   A2

DFVRF<u>TYSAGIVQI</u>LLDFVRFMGV<u>EGGVGWRH</u>W*HHHHHH*
 B44   A24   A2   B44

Figure 3B

MEDPGFFNVEIPEFPFYPTCNVCTADVNVTINFDVGGKKHQLDLDF<u>GQLTPHTKAVYQPRGAFGGSENAT</u>
<div align="right">T-help epitope   70</div>

NLFLLELLGAGELALTMRSKKLPINVTTGEEQQVSLESVDVYFQDVFGTMWCHHAEMQNPVYLIPETVPY
<div align="right">140</div>

IKW<u>DNCNSTNITAVVRAQGLD</u>VTLPLSLPTSAQDSNFSVKTEMLGNEIDIECIMEDGEISQVLPGDNKFN
    T-help epitope

ITCSGYESHVPSGGILTSTSPVATPIPGTGYAYSLRLTPRPVSRFLGNNSILYVFYSGNGPKASGGDYCI
<div align="right">210</div>

QSNIVFSDEIPASQDMPTNTTDITYVGDNATYSVPMVTSEDANSPNVTVTAFWAWPNNTETDFKCKWTLT
<div align="right">280</div>

SGTPSGCENISGAFASNRTFDITVSGLGTAPKTLIITRTATNATTTTHKVIFSKAPESTTTSPTLNTTGF
<div align="right">350</div>

ADPNTTTG*VDF*<u>LRGRAYGL</u>PYLFW<u>LAAI</u><u>HRCQAIRKKRRIYDLIEL</u>G<u>LCTLVAML</u><u>RLRAE</u>AQVK<u>IEDPPF</u>
<div align="right">420</div>

  428   B8   A23   B27   B27   A2   A3   B60

<u>NSL</u>TYSAGIVQILLDFVRFMGV<u>EGGVGWRH</u>WIA<u>LYLQQNWWT</u>L<u>RAKFKQLL</u>SSCSSCPLSKITYGPVFMC
 A24   A2   B44   B57/A2   B8   A11   A24

<u>LQAKWRLQTL</u><u>RPPIFIRRL</u><u>VSFIEFV</u>GWY<u>PLHEQHGM</u>VEITPYKPTW<u>CLGGLLTMV</u>EENLLDFVRFYLLE
  B8   B7   B58   B35   B44   A2   B44

MLWR<u>LGQGGSPTAMAVLLHEESM</u>VQPPQLTLQV*HHHHHH*
 A2   B62   B35   B46

Figure 4A

Figure 4B

Figure 4C

Figure 5

gp350PT26A-ISCOM vaccine (sub. cut)

gp350PT26B-ISCOM vaccine (sub. cut)

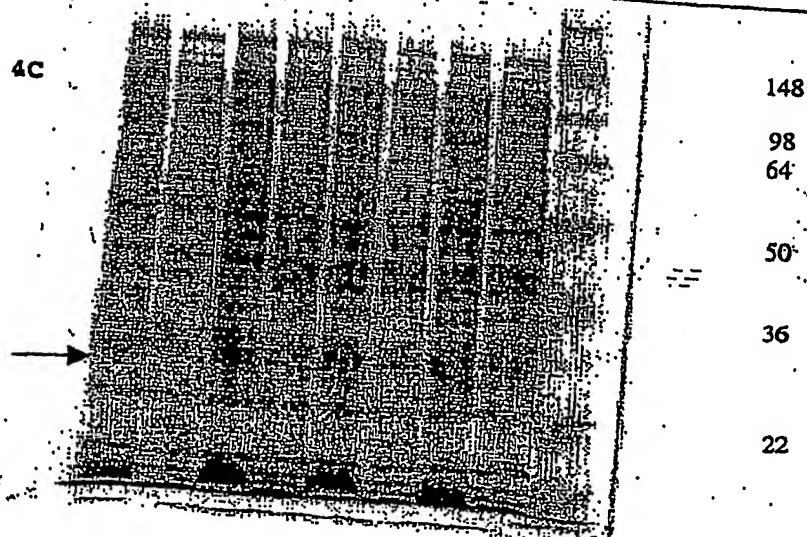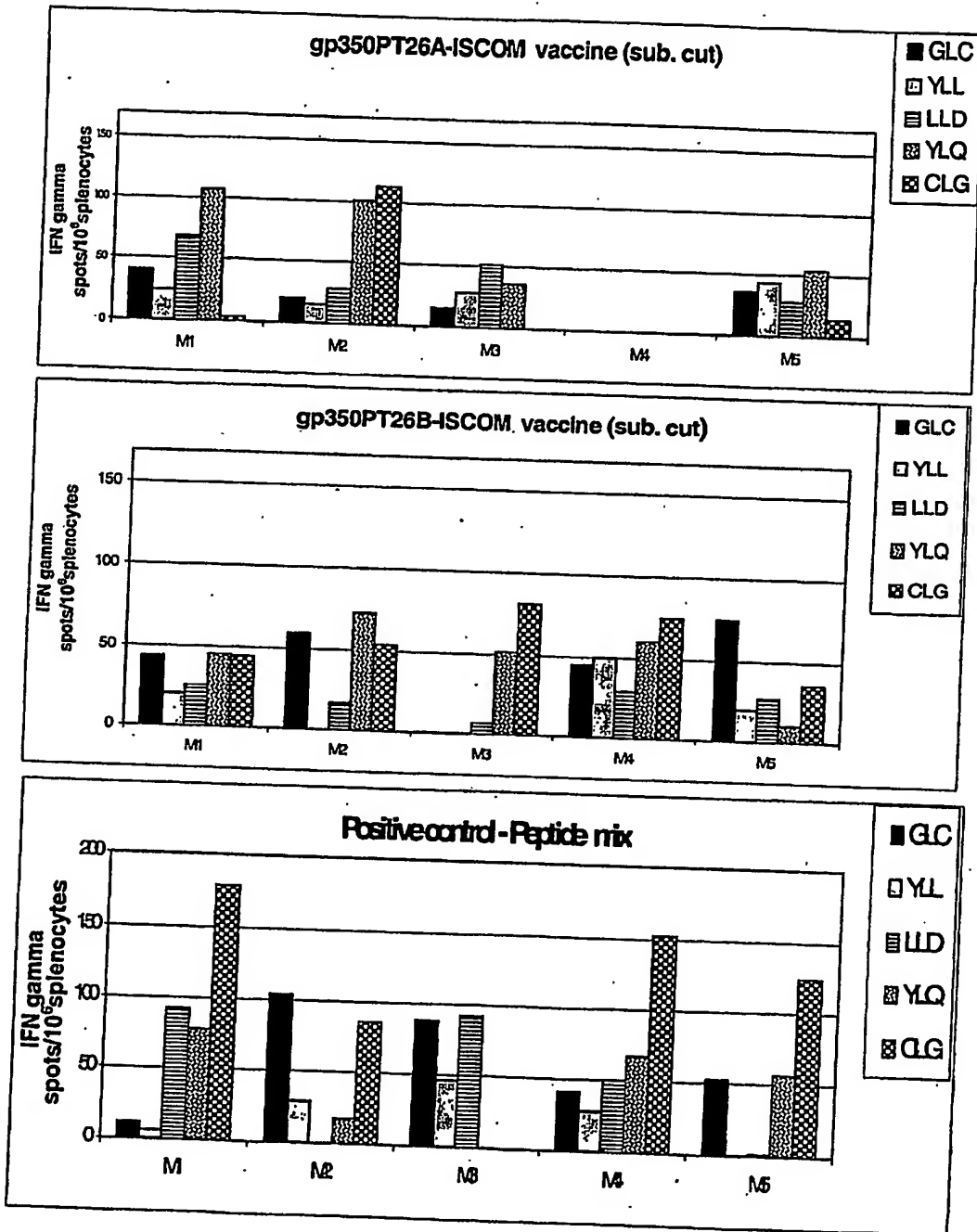Positive control - Peptide mix

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/AU03/000910

International filing date: 14 July 2003 (14.07.2003)

Document type: Certified copy of priority document

Document details: Country/Office: AU
Number: 2002950183
Filing date: 12 July 2002 (12.07.2002)

Date of receipt at the International Bureau: 05 September 2005 (05.09.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse